

基于改进混合高斯模型的运动目标检测算法

孙毅刚, 李 鸿, 张红颖

(中国民航大学航空自动化学院, 天津 300300)

摘 要: 在运动目标停滞的情况下, 传统的混合高斯模型会将运动目标误判为背景, 导致目标漏检。为此, 提出一种基于改进混合高斯模型的目标检测算法。引入背景学习参数, 结合前一帧的检测结果自适应地更新背景, 从而提取完整的运动目标。利用像素的八连通区域信息抑制噪声, 提高算法在复杂环境中的稳定性。实验结果表明, 与传统检测方法相比, 该算法能够在复杂环境中准确地检测出短暂停滞的运动目标。

关键词: 混合高斯模型; 运动目标检测; 背景模型; 噪声抑制; 连通区域

Moving Object Detection Algorithm Based on Improved Gaussian Mixture Model

SUN Yi-gang, LI Hong, ZHANG Hong-ying

(College of Aeronautical Automation, Civil Aviation University of China, Tianjin 300300, China)

【Abstract】 Moving objects can be converted into background ones by Gaussian Mixture Model(GMM) when they might be staying still in the scene for an uncertain time. Therefore, a new method based on improved GMM is proposed. To obtain complete objects, a background learning parameter is introduced to update the model according to the detection result in previous frame. Moreover, the information of 8-adjacent connection area is utilized to suppress noises and improve its stability in the complex environment. Several experiments are implemented and the results demonstrate its effectiveness in detecting the moving objects which stay briefly in the complicated condition.

【Key words】 Gaussian Mixture Model(GMM); moving object detection; background model; noise suppression; connection area

DOI: 10.3969/j.issn.1000-3428.2012.18.045

1 概述

运动目标检测是视频图像分析的重要部分, 也是对视频做进一步处理(编码、目标跟踪、目标分类、目标行为理解等)的基础。运动目标检测方法主要有光流算法、帧间差分法和背景减除法3种, 其中, 背景减除法是目前研究最多、应用最广泛的方法。文献[1]提出的自适应混合高斯模型(Gaussian Mixture Model, GMM)为每个像素建立 K 个高斯模型, 很好地描述了复杂背景, 在实时监控和目标跟踪方面能发挥很大的优势。文献[2]在自适应混合模型的基础上通过为每个像素计算自适应学习率以增加算法的收敛速度。文献[3]提出基于像素点的自适应密度估计的高斯混合模型, 用背景相减实现目标检测。文献[4]通过改进学习率和方差的更新方式提高了目标检测的速度。文献[5]通过亮度归一化和不同像素点的特点对图像进行分类处理, 减少了混合高斯模型的运算量。文献[6]通过核函数估计并结合最大后验概率马尔可夫随机场提高了混合高斯模型的自适应性。文献[7]通过增量式主成分分析并结合三帧差分法实现了对运动目标的提取。文献[8]在颜色

和空间上对像素进行建模, 减少了环境的干扰, 增强了混合高斯模型的自适应性。然而, 这些方法对运动速度变化的物体检测效果较差, 特别是当运动物体在场景中停滞一段时间, 会造成漏检。

针对该情况, 本文通过在混合高斯模型中引入背景学习参数, 改进了背景模型更新方式, 并且结合像素的八连通区域, 能够在复杂场景中完整地提取出运动目标, 避免“空洞”现象。

2 混合高斯背景建模

混合高斯模型^[1]对视频图像中的每一像素点用 K 个高斯模型去近似(K 一般取3~5之间), 用新得到的观察值更新模型参数, 使旧观察值的影响力逐渐变小。若每个像素点颜色取值用变量 X_i 表示, 其概率密度函数可用如下 K 个三维高斯函数表示:

$$p(X_t) = \sum_{i=1}^K w_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

其中, $\eta(X_t, \mu_{i,t}, \Sigma_{i,t})$ 是 t 时刻的第 i 个高斯分布; $\mu_{i,t}$ 和 $\Sigma_{i,t}$ 分别为高斯分布的均值和方差, $\Sigma_{i,t} = \sigma_{i,t}^2 I$ (假设像素点的

基金项目: 国家自然科学基金委员会与中国民用航空局联合基金资助项目(60979005); 中央高校基本科研业务费专项基金资助项目(ZXH2009B004); 中国民航局科技基金资助项目(MHRD201002)

作者简介: 孙毅刚(1963—), 男, 教授、博士, 主研方向: 目标监控, 图像处理; 李 鸿, 硕士研究生; 张红颖, 副教授、博士

收稿日期: 2011-11-01 **修回日期:** 2012-01-04 **E-mail:** ygsun@cauc.edu.cn

RGB 各分量互相独立, 且有相同的方差, I 为单位矩阵); $w_{i,t}$ 是第 i 个高斯分布的权重, 且有 $\sum_{i=1}^K w_{i,t} = 1$ 。其中:

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)}, i = 1, 2, \dots, K \quad (2)$$

其中, n 表示 X_t 的维数。

对新获取的像素值 X_t , 判断其与已存在的 K 个高斯分布是否匹配, 若满足条件 $|X_t - \mu_{i,t-1}| \leq 2.5\sigma_i$, 则认为该像素值与高斯分布匹配。如果 X_t 都不匹配, 则增加新高斯分布或用新高斯分布代替优先级 $\lambda_{i,t} = w_{i,t}/\sigma_{i,t}$ ($i = 1, 2, \dots, K$) 最小的分布, 新的高斯分布以 X_t 作为均值。

高斯分布权重、均值和方差分别按式(3)~式(5)更新。对于匹配的分量, $M_{k,t} = 1$, 否则, $M_{k,t} = 0$ 。对于不匹配的分量, 均值和方差都保持不变。

$$w_{i,t} = (1 - \alpha)w_{i,t-1} + \alpha(M_{i,t}) \quad (3)$$

$$\mu_{i,t} = (1 - \rho)\mu_{i,t-1} + \rho X_{i,t} \quad (4)$$

$$\sigma_{i,t}^2 = (1 - \rho)\sigma_{i,t-1}^2 + \rho(X_{i,t} - \mu_{i,t})^T (X_{i,t} - \mu_{i,t}) \quad (5)$$

$$\rho = \alpha \eta(X_t, \mu, \Sigma) \quad (6)$$

其中, α 是用户定义的学习率, 且 $0 \leq \alpha \leq 1$, α 值决定了背景更新速度。高斯模型更新完毕后, 重新按优先级 $\lambda_{i,t}$ 排列高斯模型, 取前 B 个高斯分布生成背景:

$$B = \arg\min_b \left(\sum_{K=1}^B w_K > T \right) \quad (7)$$

其中, T 是一个背景阈值, 用于确定作为背景模型的分布。

3 本文的运动目标检测算法

由于传统混合高斯背景模型采用固定更新率, 难以适应目标运动速度变化的情况, 对于大而慢的运动目标检测效果差, 只能检测到目标小部分轮廓。特别是当运动目标在场景中停滞一段时间(如刚停在停车场中的汽车)时, 传统的混合高斯模型很可能将运动目标误判为背景, 产生“空洞”现象。

虽然文献[9]为每个像素引入参考量, 提高了背景模型的自适应性, 同时结合基于区域的像素检测方法使上述情况有一定的改善, 但是该方法在复杂场景中会产生噪声, 而且对环境变化的适应性较差。因此, 本文对传统的混合高斯模型进行了改进, 以提高混合高斯模型在复杂场景中对运动目标检测的准确性。

3.1 改进的混合高斯模型

当运动目标由运动转为静止时, 传统的混合高斯模型不断用新得到的像素观测值更新模型参数, 这使得检测到的运动目标前景逐渐变淡直至消失, 导致运动目标的漏检。为避免上述情况发生, 当运动目标静止时, 应更改背景模型更新方式。因此, 在混合高斯模型中引入背景学习参数, 定义为:

$$Train_{x,t} = (1 - \alpha \times M) \times Train_{x,t-1} + \alpha \times M \times X_t \quad (8)$$

其中, $Train_{x,t}$ 是在 t 帧当前像素 X_t 引入的一个变量; X_t 是当前像素观测值;

$$M = \begin{cases} 1 & \text{若 } X_{t-1} \text{ 为背景点} \\ 0 & \text{若 } X_{t-1} \text{ 为前景点} \end{cases} \quad (9)$$

$Train_{x,t}$ 初始值为第 1 帧序列图像的像素值 X_1 , 即 $Train_{x,1} = X_1$ 。当 $t-1$ 帧像素点 X_{t-1} 为背景点时, 则 $Train_{x,t} = (1 - \alpha)Train_{x,t-1} + \alpha X_t$, $Train_{x,t}$ 对背景实时更新。当 $t-1$ 帧像素点 X_{t-1} 为前景点时, $Train_{x,t} = Train_{x,t-1}$, $Train_{x,t}$ 一直记录着像素 X 被判为前景之前的像素观测值(即作为背景时的像素观测值)。

加入约束条件: $|X_t - Train_{x,t}| \leq T_1$ (T_1 是用户自定义的阈值, 决定算法的精度)。若当前像素值 X_t 与变量 $Train_{x,t}$ 关系满足该条件, 则说明当前像素观测值与背景相近, 用混合高斯模型对其建模, 进行进一步判断; 若当前像素不满足该条件, 则直接判断该像素为前景, 无需进行混合高斯建模。当运动目标由运动转为静止, $Train_{x,t}$ 一直记录着运动目标静止前的像素观测值, 使当前像素 X_t 不满足约束条件, 即像素 X_t 被判断为前景点, 而且不会对该像素进行混合高斯建模, 因此, 该像素不会因运动目标停滞时间太长被误认为背景点, 形成“空洞”, 从而避免了运动目标的漏检。

3.2 改进的前景检测

经上述处理后, 仍然会有一些由光照改变、树枝摆动等因素产生的噪声。在检测结果中这些噪声一般为孤立点或条纹, 而属于运动目标的像素区域具有连通性。由此对前景检测提出一种改进算法, 即结合当前像素的八连通区域对当前像素进行判断, 如图 1 所示。

E	A	H
E	m	D
F	C	G

图 1 m 的八连通区域

在图 1 中, 在当前像素 m 为前景时, 若其八连通区域在前后 2 帧的图像中都为背景点, 则可认为像素点 m 为噪声, 应判断为背景点, 否则, 需要对像素点 m 的四连通区域做进一步判断: 在四连通区域 ABCD 中, 若背景点的个数大于前景点的个数, 则像素点 m 为背景点; 反之, 像素点 m 为前景点。若背景点的个数与前景点相等, 则会出现如图 2 所示的情况(其中, 白色区域表示前景点, 灰色区域表示背景点)。此时需要判断八连通区域 EFGH 为前景点的个数是否大于等于 2, 若成立, 则像素点 m 为前景点, 否则为背景点。如果八连通区域 EFGH 为前景点的个数等于 2, 则出现如图 3、图 4 所示情况。

E	A	H
B	m	D
F	C	G

(a) 第 1 种情况

E	A	H
B	m	D
F	C	G

(b) 第 2 种情况

图 2 区域 ABCD 为背景点的个数等于前景点个数

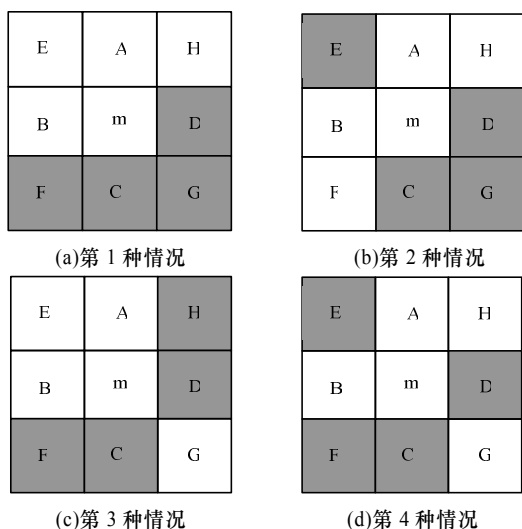


图 3 区域 EFGH 为前景点的个数等于 2 时图 2(a)所对应的情况

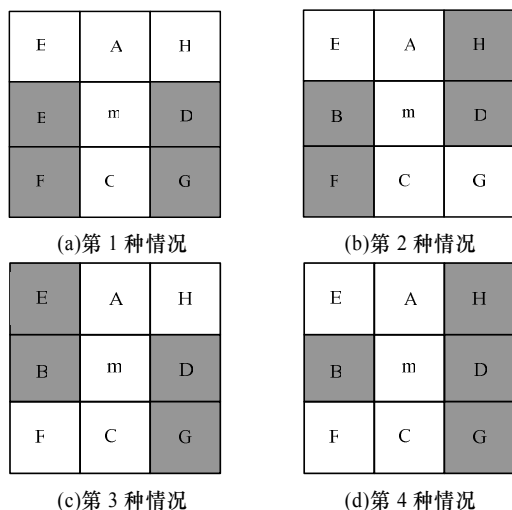


图 4 区域 EFGH 为前景点的个数等于 2 时图 2(b)所对应的情况

由于上述情况中前景像素点的区域都具有连通性,因此像素点 m 可作为运动目标,被判断为前景点。同理,当像素 $EFGH$ 为前景点的个数大于 2 时连通性更普遍,像素点 m 为前景点的可能性更大。

4 实验结果

实验条件为: Pentium Dual-Core 2.70 GHz CPU, 2 GB 内存, Visual Studio 2008 开发平台, 运用 OpenCV2.2 进行编程实现。

依照本文算法分别对 320×240 和 432×324 2 段视频序列图像进行了背景建模, 速度分别为 27 f/s 和 25 f/s, 满足实时检测要求。取背景学习率 $\alpha = 0.005$, 混合高斯模型的个数 $K = 5$, $T_1 = 25$, 初始化标准差 $\sigma = 30$ 。

第 1 段视频是在室内拍摄, 如图 5 所示。在视频序列图像中, 分别截取第 107 帧、第 314 帧、第 366 帧图像进行处理。其中, 第 107 帧图像中运动目标(人)正缓慢前进, 而第 314 帧图像中运动目标突然静止, 在运动目标静止一段时间后获取第 366 帧图像。图 6~图 8 分别是用传统混合高斯模型、文献[9]方法和本文方法对视频序列图像的处理结果。



图 5 原始视频序列 1

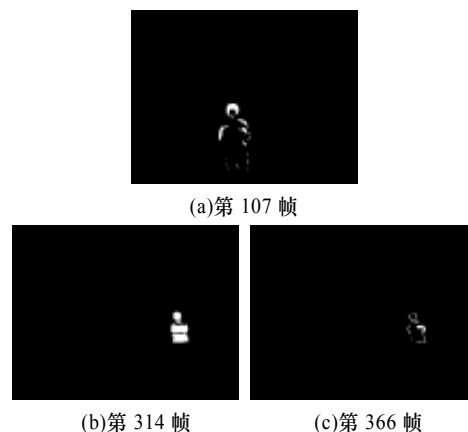


图 6 传统混合高斯模型对原始视频序列 1 的处理结果

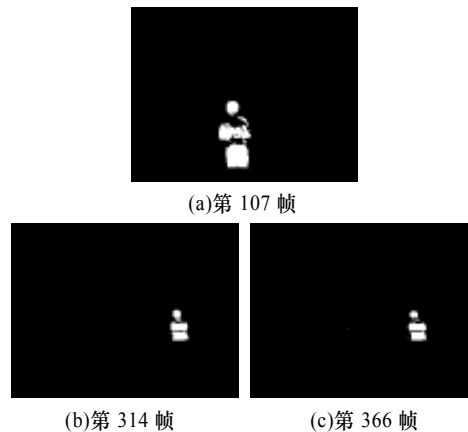


图 7 文献[9]方法对原始视频序列 1 的处理结果

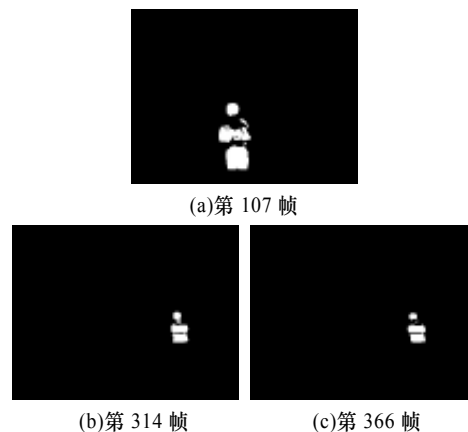


图 8 本文方法对原始视频序列 1 的处理结果

由图 6(a)与图 7(a)、图 8(a)的对比可以看出, 传统混合高斯模型对场景中大而慢的运动目标检测效果差, 只能检测到目标轮廓, 而文献[9]方法和本文方法可以检测到目标的整体。由图 6 可以看出, 当运动目标突然静止, 传统混合高斯模型能检测到运动目标, 然而随着时间的推移, 检测到的运动目标逐渐变淡, 形成“空洞”, 如图 6(c)所示。若运动目标继续保持静止, 传统混合高斯模型会把运动目标误认为背景, 导致目标的漏检。而对于文献[9]方法和本文方法, 即使运动目标静止一段时间, 也会一直保持对运动目标的检测, 不会产生“空洞”现象, 如图 7(c)和图 8(c)所示。

第 2 段视频在有树枝摆动的复杂环境中拍摄, 如图 9 所示。在视频序列图像中, 分别截取第 32 帧、第 252 帧、第 391 帧图像进行处理。第 32 帧是运动目标(汽车)进入场景之前的截图, 第 252 帧是运动目标(汽车)在场景中停滞一段时间的截图, 第 391 帧是运动目标(汽车)即将离开场景时的截图。图 10~图 12 分别是用传统混合高斯模型、文献[9]方法方法和本文改进方法对视频序列图像的处理结果。



图 9 原始视频序列 2

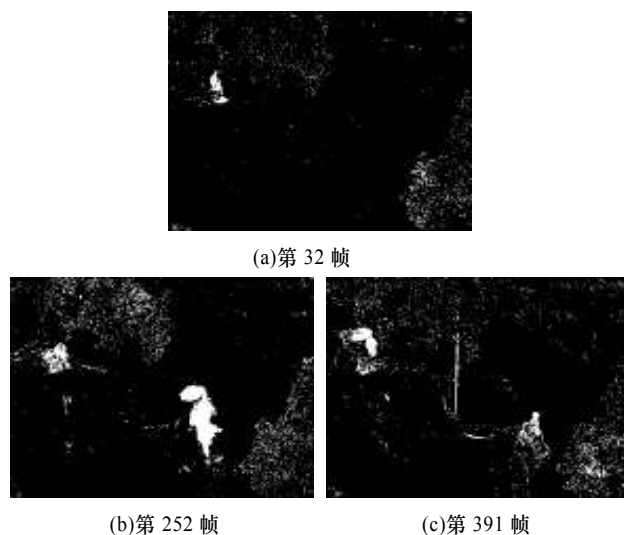


图 10 传统混合高斯模型对原始视频序列 2 的处理结果

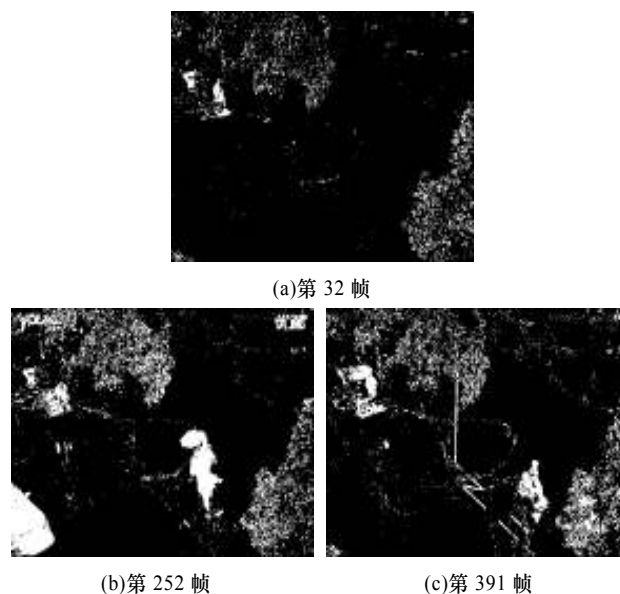


图 11 文献[9]方法对原始视频序列 2 的处理结果

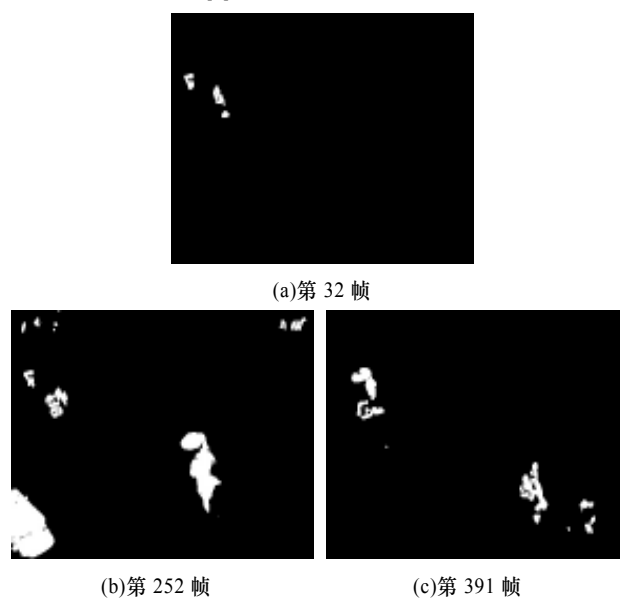


图 12 本文方法对原始视频序列 2 的处理结果

通过比较图 10、图 11 与图 12 可以看出: 传统的混合高斯模型在复杂环境中对噪声的处理效果差, 而且当运动目标(汽车)在场景中静止一段时间时, 会将运动目标误认为背景, 导致漏检, 如图 10(b)所示; 文献[9]的方法虽然能够完整地检测出运动目标, 不会产生漏检现象, 但是在复杂场景中产生了噪声, 影响对运动目标的检测, 如图 11 所示; 而本文算法可以将目标完整地提取出来, 并且可以有效地抑制噪声, 如图 12 所示。

为了得到定量的误差分析, 对第 2 段视频手动生成前景目标, 并以此作为标准, 采用消极误检率(False Negative Rate, FNR)和积极误检率(False Positive Rate, FPR)分析实验结果, 如表 1 所示。消极误检率反映了算法对目标的漏检问题, 而积极误检率反映了算法受环境噪声的影响程度。由此可以看出, 本文算法的消极误检率比传统混合高斯模型算法小, 避免了对运动目标的漏检, 同时积极误检率明显比后者小, 从而减小了噪声的影响。

表1 本文算法与GMM算法的误检率

实验 组别	FNR		FPR	
	本文算法	GMM 算法	本文算法	GMM 算法
1	0.010 1	0.011 8	0.002 1	0.214 9
2	0.008 9	0.020 3	0.004 3	0.294 6
3	0.023 1	0.086 5	0.003 5	0.154 1
4	0.043 9	0.135 0	0.007 8	0.378 1
5	0.056 1	0.121 0	0.011 3	0.438 0
6	0.049 9	0.141 3	0.010 2	0.397 7
7	0.034 6	0.095 3	0.004 6	0.356 1
8	0.028 9	0.078 0	0.006 1	0.286 3
9	0.014 5	0.013 4	0.009 3	0.312 0
10	0.023 1	0.021 6	0.007 4	0.133 0

5 结束语

针对运动目标移动缓慢或静止的情况,本文通过引入背景学习参数并结合像素的八连通区域对混合高斯模型进行了改进,能完整地检测出短暂停滞的运动目标,抑制复杂环境变化产生的噪声。实验结果证明,本文的改进算法能够有效地提高检测准确性。

参考文献

- [1] Stauffer C, Grimson W. Adaptive Background Mixture Models for Real-time Tracking[C]//Proc. of IEEE Conference on Computer Vision and Pattern Recognition. Fort Collins, USA: IEEE Press, 1999: 246-252.
- [2] Lee Dar-Shyang. Effective Gaussian Mixture Learning for Video Background Subtraction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(5): 827-832.
- [3] Zivkovic Z, van der Heijden F. Efficient Adaptive Density Estimation per Image Pixel for the Task of Background Subtraction[J]. Pattern Recognition Letters, 2006, 27(7): 773-780.
- [4] Bouttefroy P L M, Bouzerdoum A, Phung S L, et al. On the Analysis of Background Subtraction Techniques Using Gaussian Mixture Models[C]//Proc. of IEEE International Conference on Acoustics Speech and Signal Processing. Dallas, USA: IEEE Press, 2010.
- [5] 徐以美, 郭宝龙, 张 晋. 基于像素分类的运动目标检测算法[J]. 计算机工程, 2008, 34(23): 205-226.
- [6] 周箴毅, 胡福乔. 基于背景建模的动态场景目标检测[J]. 计算机工程, 2008, 34(24): 203-205.
- [7] 牛武泽, 石林锁, 金广智, 等. 融合时空信息的运动目标检测算法[J]. 计算机工程, 2011, 37(18): 171-176.
- [8] Dickinson P, Hunter A, Appiah K. A Spatially Distributed Model for Foreground Segmentation[J]. Image and Vision Computing, 2009, 27(9): 1326-1335.
- [9] Huang Tianci, Qiu Jingbang, Sakayori T. Motion Detection Based on Background Modeling and Performance Analysis for Outdoor Surveillance[C]//Proc. of IEEE International Conference on Computer Modeling and Simulation. Macau, China: [s. n.], 2009.
- [10] Kok S, Domingos P. Learning the Structure of Markov Logic Networks[C]//Proc. of the 22nd Int'l Conf. on Machine Learning. Bonn, Germany: AAAI Press, 2005: 441-448.
- [11] Mihalokva L, Mooney R J. Bottom-up Learning of Markov Logic Network Structure[C]//Proc. of the 24th Int'l Conf. on Machine Learning. Bonn, Germany: AAAI Press, 2007: 625-632.
- [12] Singla P, Domingos P. Discriminative Training of Markov Logic Networks[C]//Proc. of the 20th National Conf. on Artificial Intelligence. Pittsburgh, USA: AAAI Press, 2005: 868-873.
- [13] Liu D C, Nocedal J. On the Limited Memory BFGS Method for Large Scale Optimization[J]. Mathematical Programming, 1989, 45(1-3): 503-528.
- [14] Singla P, Domingos P. Memory-efficient Inference in Relation Domains[C]//Proc. of the 21st National Conf. on Artificial Intelligence. Boston, USA: AAAI Press, 2006: 488-493.
- [15] Gilks W R, Richardson S, Spiegelhalter D J. Markov Chain Monte Carlo in Practice[M]. London, UK: Chapman and Hall, 1996.
- [16] Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by Simulate Annealing[J]. Science, 1983, 220(4598): 671-680.
- [17] Poon H, Domingos P. Sound and Efficient Inference with Probabilistic and Deterministic Dependencies[C]//Proc. of the 21st National Conf. on Artificial Intelligence. Boston, USA: AAAI Press, 2006: 458-463.
- [18] Neal R M. Slice Sampling[J]. The Annals of Statistics, 2003, 31(3): 705-767.
- [19] Banko M, Cafarella M J. Open Information Extraction from theWeb[C]//Proc. of IJCAI'09. New York, USA: AAAI Press, 2009.
- [20] 廖志华. 马尔科夫逻辑网络在引文匹配和中文命名实体识别中的应用研究[D]. 重庆: 西南大学, 2009.
- [21] Kok S, Domingos P. Extracting Product Information from Email Receipts Using Markov Logic[C]//Proc. of the 6th Conf. on Email and Anti-spam. Cambridge, USA: MIT Press, 2009.
- [22] Banko M, Etzioni O. The Tradeoffs Between Open and Traditional Relation Extraction[C]//Proc. of the 46th Annual Meeting of the Association for Computational Linguistics. Cambridge, USA: MIT Press, 2008: 28-36.
- [23] Zhu Jun, Zaiqing N. StatSnowball: A Statistical Approach to Extracting Entity Relationships[C]//Proc. of WWW'09. Madrid, Spain: ACM Press, 2009: 101-110.
- [24] Singla P, Domingos P. Entity Resolution with Markov Logic[C]//Proc. of the 6th IEEE International Conference on Data Mining. Hong Kong, China: IEEE Press, 2006: 572-582.
- [25] 楼俊杰, 徐从富, 郝春亮. 基于马尔科夫逻辑网络的实体解析改进算法[J]. 计算机科学, 2010, 37(8): 243-247.

编辑 张正兴

编辑 张正兴

(上接第 165 页)